# Intermittent GPS-aided VIO: Online Initialization and Calibration

Woosik Lee, Kevin Eckenhoff, Patrick Geneva, and Guoquan Huang

Abstract— In this paper, we present an efficient and robust GPS-aided visual inertial odometry (GPS-VIO) system that fuses IMU-camera data with intermittent GPS measurements. To perform sensor fusion, spatiotemporal sensor calibration and initialization of the transform between the sensor reference frames are required. We propose an online calibration method for both the GPS-IMU extrinsics and time offset as well as a reference frame initialization procedure that is robust to GPS sensor noise. In addition, we prove the existence of four unobservable directions of the GPS-VIO system when estimating in the VIO reference frame, and advocate a state transformation to the GPS reference frame for full observability. We extensively evaluate the proposed approach in Monte-Carlo simulations where we investigate the system’s robustness to different levels of GPS noise and loss of GPS signal, and additionally study the hyper-parameters used in the initialization procedure. Finally, the proposed system is validated in a large-scale real-world experiment

摘要：在本文中，我们提出了一种高效且强大的 GPS 辅助视觉惯性里程计 (GPS-VIO) 系统，该系统将 IMU 相机数据与间歇性 GPS 测量相融合。为了执行传感器融合，需要时间空间传感器校准和传感器参考坐标系之间的变换初始化。

我们提出了一种针对 GPS-IMU 外在参数和时间偏移的在线校准方法，以及对 GPS 传感器噪声具有鲁棒性的参考坐标系初始化程序。此外，我们证明了在 VIO 参考系中进行估计时 GPS-VIO 系统存在四个不可观测的方向，并提倡对 GPS 参考框架进行状态转换以实现完全可观测性。我们在 Monte-Carlo 模拟中广泛评估了所提出的方法，我们研究了系统对不同级别的 GPS 噪声和 GPS 信号丢失的鲁棒性，并另外研究了初始化过程中使用的超参数。最后，所提出的系统在大规模的真实世界实验中得到验证

1. INTRODUCTION AND RELATED WORK

For any autonomous robotic system, robust and accurate localization is a primary requirement. Localization is typically performed by estimating the robot’s state using measurements from on-board sensors. Of many possible sensor deployments, cameras and inertial measurement units (IMUs) – which measure linear accelerations and angular velocities of the moving robot – are commonly used for 3D navigation [1] in both indoor and outdoor environments, as they are low-cost yet provide high-quality ego-motion estimation [2]–[5]. However, when only using these sensors it is difficult to provide long-term, drift-free estimation due to the accumulation of relative motion errors. A commonly used approach to bound navigation error is a simultaneous localization and mapping (SLAM) that exploits loop-closure constraints to correct the accumulated error [6], [7]. However, such methods have a major drawback of both increased computational complexity and memory requirements

对于任何自主机器人系统，稳健和准确的定位是首要要求。定位通常是通过使用来自车载传感器的测量来估计机器人的状态来执行的。在许多可能的传感器部署中，相机和惯性测量单元 (IMU) - 测量移动机器人的线性加速度和角速度 - 通常用于室内和室外环境中的 3D 导航 [1]，因为它们成本低但提供高质量的自我运动估计[2]-[5]。然而，当仅使用这些传感器时，由于相对运动误差的累积，很难提供长期、无漂移的估计。一种常用的绑定导航误差方法是同时定位和映射 (SLAM)，它利用闭环约束来纠正累积误差 [6]、[7]。然而，这种方法的一个主要缺点是计算复杂度和内存需求都增加了

As compared to SLAM, global measurement sensors, such as those from Global Positioning System (GPS), directly provide absolute position information to reduce drift. However, the accuracy of GPS measurements is highly dependent on the surrounding environment and the availability of external correction data. Synchronous sensors have been particularly considered in prior works, of which many fused inertial and GPS readings [8]–[11], with others leveraging camera, inertial and GPS sensors fusion [12]–[17] with great success. The asynchronous inclusion of GPS measurements within a sensor fusion framework remains challenging due to their low rate, high noise, and intermittency

与 SLAM 相比，全球测量传感器，例如来自全球定位系统 (GPS) 的传感器，直接提供绝对位置信息以减少漂移。然而，GPS 测量的准确性在很大程度上取决于周围环境和外部校正数据的可用性。同步传感器在之前的工作中被特别考虑，其中许多融合惯性和 GPS 读数 [8]-[11]，其他利用相机、惯性和 GPS 传感器融合 [12]-[17] 取得了巨大成功。由于其低速率，高噪音和间歇性在传感器融合框架中异步包含 GPS 测量仍然具有挑战性，

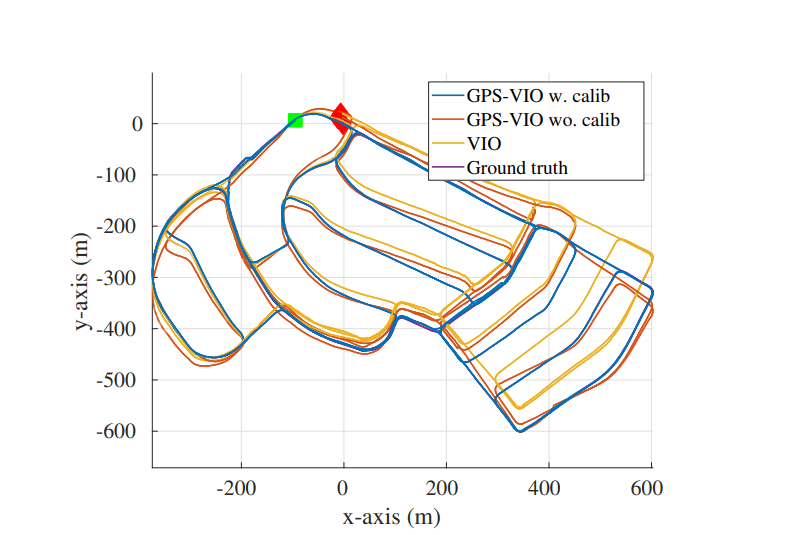


Fig. 1: Simulation results of GPS-VIO with calibration (blue), GPS-VIO without calibration (red), and VIO (yellow). The green and red squares correspond to the start and end of the 9.1km trajectory, respectively. sensor fusion framework remains challenging due to their low rate, high noise, and intermittency.

图 1：带校准的 GPS-VIO（蓝色）、不带校准的 GPS-VIO（红色）和 VIO（黄色）的仿真结果。绿色和红色方块分别对应 9.1 公里轨迹的起点和终点。传感器融合框架由于其低速率、高噪声和间歇性而仍然具有挑战性。

In order to optimally fuse multiple sensor measurements from different sensor frames, the transformation between the sensor frames and the time offset between the sensors must be known. An initial imperfect guess of the calibration between the sensor frames is often known beforehand, but if it is treated as perfect the state estimation can suffer, thus its refinement during online estimation is highly desirable. For a camera and IMU pair, the calibration of spatial and/or temporal parameters is well studied in [18]–[21]. While offline calibration of camera, IMU and GPS is often performed within an optimization framework [22], [23], online estimation of the transformation among the sensors was also investigated within a Kalman filter (KF) framework [13], [24]–[26]. However, to the best of our knowledge, no work to date has considered the estimation of the time offset between the GPS and IMU/camera, and their inherent asynchronous nature can greatly impact the estimation performance if ignored

为了最佳地融合来自不同传感器坐标系的多个传感器测量值，必须知道传感器坐标系之间的转换和传感器之间的时间偏移。传感器坐标系之间校准的初始不完美猜测通常是事先知道的，但如果它被视为完美状态估计可能会受到影响，因此在线估计期间的细化是非常可取的。对于相机和 IMU 对空间和时间参数的校准在 [18]-[21] 中得到了很好的研究。在相机离线校准时，IMU 和 GPS 通常在优化框架 [22]、[23] 内执行，传感器之间转换的在线估计也在卡尔曼滤波器 (KF) 框架内进行了研究 [13]、[24]-[26] .然而，据我们所知，迄今为止没有工作考虑过 GPS 和 IMU/相机之间的时间偏移的估计，如果忽略它们固有的异步性质会极大地影响估计性能

GPS provides latitude, longitude, and altitude readings in a geodetic coordinate frame, which is commonly converted to Cartesian coordinate East-North-Up (ENU) {E}, e.g., by setting the first GPS measurement as the datum. Conversely, a visual inertial odometry (VIO) system estimates its state relative to a starting VIO frame {V }, and is known to have four unobservable directions corresponding to the global position and yaw [27], [28]. In order to fuse GPS measurements in the global frame {E} with VIO state estimates in the {V } frame, the transformation between them must be computed, which is the “reference frame initialization” problem. Unlike sensor calibration, an initialization procedure is required to find this unknown transformation that varies from run-to-run.

GPS 在大地坐标系中提供纬度、经度和高度读数，通常将其转换为笛卡尔坐标东北天(ENU) {E}，例如，通过将第一个 GPS 测量值设置为基准。相反，视觉惯性里程计 (VIO) 系统估计其相对于起始 VIO 帧 {V} 的状态，并且已知有四个不可观察的方向对应于全局位置和偏航[27]，[28]。为了将全局帧 {E} 中的 GPS 测量值与 {V} 帧中的 VIO 状态估计融合，必须计算它们之间的转换，这就是“参考帧初始化”问题。与传感器校准不同，需要一个初始化过程来找到这个因运行而异的未知转换。

This initialization problem can be formulated as a general 3D position trajectory alignment problem. For example, Horn [29] used singular value decomposition (SVD) of a covariance matrix to derive a closed-form solution. Shepard et al. [30] leveraged this method to compute a 7 degree-offreedom (d.o.f) transformation between synchronized GPS and VIO trajectories. Umeyama [31] presented a method in the presence of large trajectory noises, which was used to find the transformation between two gravity aligned trajectories [32], [33]. Other works have employed additional information for initialization, including magnetic sensors [34]–[36], yaw calculation with a straight planar motion assumption [37], or a prior map constructed in the global frame [16]

这个初始化问题可以表述为一般的 3D 位置轨迹对齐问题。例如，Horn [29] 使用协方差矩阵的奇异值分解 (SVD) 来推导封闭形式的解。谢泼德等人。 [30] 利用这种方法计算了 7 个自由度（d.o.f) 同步 GPS 和 VIO 轨迹之间的转换。 Umeyama [31] 提出了一种存在大轨迹噪声的方法，用于找到两个重力对齐轨迹之间的转换 [32]、[33]。其他工作使用了额外的初始化信息，包括磁传感器 [34]-[36]，偏航计算采用直线平面运动假设 [37]，或在全局框架中构建的先验图 [16]

Note that the closest to this work is VINS-Fusion [12], which is a loosely-coupled estimator that fuses GPS measurements and VIO’s relative poses in a secondary optimization thread. While VINS-Fusion shows impressive performance in practice, the system (i) assumes synchronized measurements with perfectly known timestamps and an identity transformation between GPS and IMU, (ii) lacks support for online refinement of sensor calibration, and (iii) does not explicitly initialize the ENU to VIO frame transform while assuming that the estimates will converge in the ENU frame as more GPS measurements are collected.

请注意，最接近这项工作的是 VINS-Fusion [12]，它是一种松耦合的估计器，它在二级优化线程中融合 GPS 测量和 VIO 的相对姿势。虽然 VINS-Fusion 在实践中表现出令人印象深刻的性能，

该系统 (i) 假设具有完全已知的时间戳和 GPS 和 IMU 之间的身份转换的同步测量，(ii) 缺乏对传感器校准的在线改进的支持，

(iii) 没有明确初始化 ENU 到 VIO 帧的变换，同时假设随着收集到更多 GPS 测量值，估计将收敛在 ENU 坐标系中。

In this paper, we develop a tightly-coupled VIO system aided by intermittent GPS measurements to provide persistent global localization results, while focusing on spatiotemporal sensor calibration and state initialization. In particular, the key contributions of this work are the following

在本文中，我们开发了一个紧密耦合的 VIO 系统，借助间歇性 GPS 测量来提供持久的全球定位结果，同时专注于时空传感器校准和状态初始化。特别是，这项工作的主要贡献如下

* We propose a tightly-coupled multi-state constraint Kalman filter (MSCKF)-based [38] estimator to optimally fuse inertial, camera, and asynchronous GPS measurements. The system can begin with VIO only (e.g., indoors) and convert the frame of reference to the ENU frame at an arbitrarily later timestep when GPS measurements become available for fusion. This ensures that the system provides seamless localization, and once global information is available the system is able to estimate in this frame of reference.
* 我们提出了一种基于紧密耦合多状态约束卡尔曼滤波器 (MSCKF) 的 [38] 估计器，以优化融合惯性、相机和异步 GPS 测量。系统只能以 VIO 开头（例如，室内），并在 GPS 测量可用于融合时在任意稍后的时间步将参考系转换为 ENU坐标系。这确保了系统提供无缝定位，并且一旦全局信息可用，系统就能够在这个参考框架中进行估计。
* To the best of our knowledge, this is the first work that models GPS-IMU time offset and performs online calibration of both the extrinsics and time offset. We also introduce a reference frame initialization procedure that is robust to high GPS noise which leverages the solution to a quadratic constraint least-squares problem [39]. We numerically analyze the choice of this procedure’s hyper-parameters under different GPS measurement noise levels
* 据我们所知，这是第一个对 GPS-IMU 时间偏移建模并执行外部和时间偏移在线校准的工作。我们还介绍了一种对高 GPS 噪声具有鲁棒性的参考帧初始化程序，该程序利用了二次约束最小二乘问题的解决方案 [39]。我们数值分析了不同 GPS 测量噪声水平下该过程的超参数的选择
* We perform an observability analysis of the GPS-VIO system to show that there are four unobservable directions if the 4 d.o.f transformation between the ENU to VIO frames is kept in the state vector, while the system is fully observable if estimating in the ENU frame
* 我们对 GPS-VIO 系统进行了可观测性分析，表明如果 ENU 到 VIO 帧之间的 4 自由度变换保持在状态向量中，则存在四个不可观测方向，而如果在 ENU 中进行估计，则系统是完全可观测的框架
* We evaluate the proposed GPS-VIO extensively in simulations, showing the calibration convergence under different measurement noise levels and the robustness to loss of GPS. Moreover, the proposed method is also validated on a real-world, large-scale experiment with both indoor and outdoor portions which exhibits varying GPS noise levels.
* 我们在模拟中广泛评估了所提出的 GPS-VIO，显示了在不同测量噪声水平下的校准收敛性和对 GPS 丢失的鲁棒性。此外，所提出的方法还在真实世界的大规模实验中得到验证，室内和室外部分表现出不同的 GPS 噪声水平。

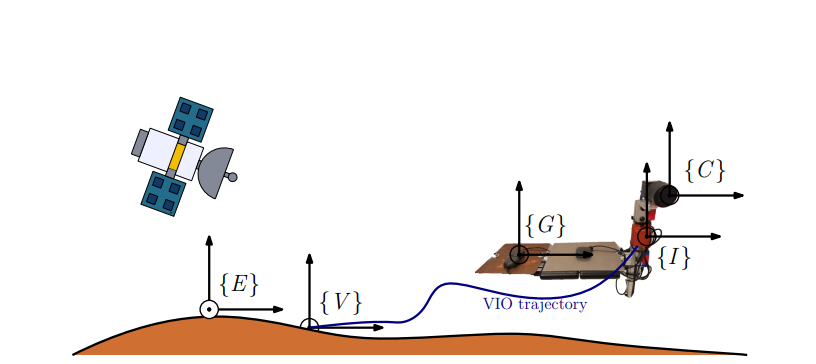


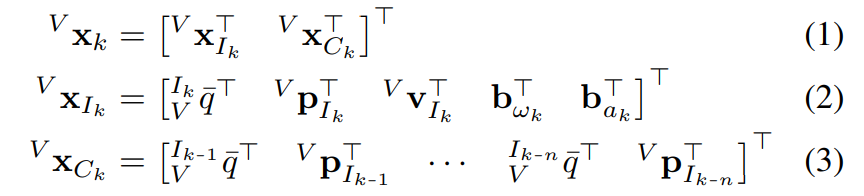
Fig. 2: Our integrated sensor system is composed of five different frames: ENU frame {E}, VIO frame {V }, IMU sensor frame {I}, camera sensor frame {C}, and GPS sensor frame {G}. {E} is the frame of the reference of the GPS measurements and {V } is the local frame set up by VIO whose orientation is aligned with gravity.

图 2：我们的集成传感器系统由五个不同的框架组成：ENU 框架 {E}、VIO 框架 {V}、IMU 传感器框架 {I}、相机传感器框架 {C} 和 GPS 传感器框架 {G}。 {E} 是 GPS 测量的参考坐标系，{V} 是 VIO 设置的局部坐标系，其方向与重力对齐。

1. PRELIMINARIES: MSCKF BASED VIO

The standard VIO state at tiestep k consists of the current inertial state and n historical IMU pose clones [38]. All states are represented in the arbitrarily chosen gravity aligned frame of reference, {V }, see Fig. 2:

时间步长 k 的标准 VIO 状态 由当前惯性状态和 n 个历史 IMU 姿势克隆 [38] 组成。所有状态都在任意选择的重力对齐参考系 {V} 中表示，见图 2：



where is the JPL unit quaternion [40] corresponding to the rotation from {V } to {I} (i.e., rotation matrix ), and are the position and velocity of {I} in {V }, and and are the gyroscope and accelerometer biases, respectively. We define , where x is the true state, is its estimate, is the error state, and the operation which maps a manifold element and its correction vector to an updated element on the same manifold [41].

其中 是 JPL 单元四元数 [40] 对应于从 {V} 到 {I} 的旋转（即旋转矩阵 ）， 和 是 {I} 在 { V } 和 和 分别是陀螺仪和加速度计偏差。我们定义 ,，其中 x 是真实状态， 是它的估计， 是错误状态，以及将流形元素及其校正向量映射到同一流形上的更新元素的操作 [41]

1. State Propagation

The linear acceleration am and angular velocity ωm measurements of the IMU are used for propagation：

IMU 的线加速度 am 和角速度 ωm 测量值用于传播



where a and ω are true acceleration and angular velocity, g ≈ [0 0 9.81]> is the global gravity, and na and nω are zero mean Gaussian noises. These measurements are used to propagate the IMU state from timestep k to k + 1 based on the following generic nonlinear kinematic model [40]:

其中 a 和 ω 是真实加速度和角速度，g ≈ [0 0 9.81]> 是全局重力，na 和 nω 是零均值高斯噪声。这些测量值用于根据以下通用非线性运动学模型 [40] 将 IMU 状态从时间步 k 传播到 k + 1：



where denotes the estimate at timestep a processing the measurements up to timestep b. We linearize (5) at the current estimate and propagate the covariance forward in time:

其中表示在时间步 a 处的估计值，处理直到时间步 b 的测量值。我们在当前估计处线性化 (5) 并及时向前传播协方差：



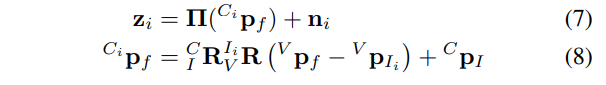
where and are the state transition matrix and discrete noise covariance [38]

其中 Φ 和 Q 是状态转移矩阵和离散噪声协方差 [38]。

1. Visual Measurement Update

We maintain a number of stochastic clones in , and perform visual feature tracking to obtain series of visual bearing measurements to 3D environmental features. A measurement at timestep is expressed as a function of a cloned pose and feature position :

我们在 ,中维护了许多随机克隆，并执行视觉特征跟踪以获得对 3D 环境特征的一系列视觉方位测量。时间步长 处的测量值 表示为克隆姿势和特征位置 的函数：



Where is the perspective projection, and and represent the camera to IMU extrinsics. By stacking all measurements for a given feature, the corresponding linearized residuals is given by:

其中是透视投影，和代表相机到IMU外参。通过堆叠给定特征的所有测量值，对应的线性化残差由下式给出：



where and are the measurement Jacobians of the state and the feature. The key idea of the MSCKF is to find the left null space of and left multiply (9) by it to infer a new measurement function that depends only on the state:

其中 和 是状态和特征的测量雅可比矩阵。 MSCKF 的关键思想是找到 的左零空间并左乘 (9) 以推断出仅取决于状态的新测量函数：



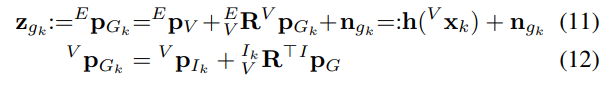
which can be directly used in an EKF update without storing features in the state, leading to substantial computational savings as the problem size remains bounded over time.

可以直接在 EKF 更新中使用，而无需在状态中存储特征，从而导致大量计算节省，因为问题大小随着时间的推移保持有界

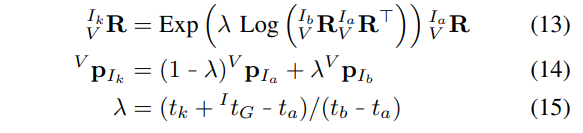
1. GPS MEASUREMENT UPDATE AND CALIBRATION

Besides the visual measurement update as in the standard MSCKF, whenever a new GPS measurement in the ENU frame {E} is available, we will update the state with it. This requires knowledge of the transform between the two frames , which we will explain in the next section. In particular, the GPS measurement at timestep k is:

除了标准 MSCKF 中的视觉测量更新之外，每当 ENU 帧 {E} 中的新 GPS 测量可用时，我们将使用它更新状态。这需要了解两帧 之间的变换，我们将在下一节中解释。尤其，时间步长 k 的 GPS 测量值 为：



Where is the GPS to IMU extrinsic calibration and is a white Gaussian noise. We note that while here is written as a full rotation matrix, we represent it as one that only rotates about the global gravity aligned z-axis. Due to the delayed asynchronous nature of the GPS sensor, the state has likely advanced beyond the collection time and thus we express the measurement as a function of the available stochastic clones. Using linear interpolation [42], the IMU pose in Eq. (12) can be expressed as:

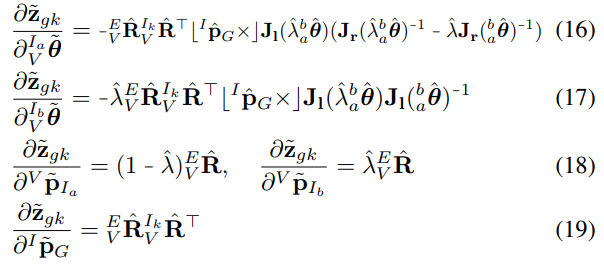
其中 GPS 到 IMU 外部校准，是高斯白噪声。我们注意到，虽然这里 被写成一个完整的旋转矩阵，但我们将其表示为仅围绕全局重力对齐的 z 轴旋转的矩阵。由于 GPS 传感器的延迟异步特性，状态可能已经超过了收集时间，因此我们将测量表示为可用随机克隆的函数。使用线性插值[42]，方程中的 IMU 姿势。 (12) 可以表示为：

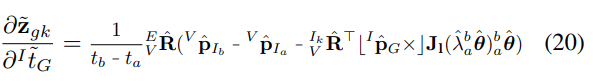
where is the time offset between the GPS and IMU clocks, the bounding poses have timestamps, and , are the matrix exponential and logarithmic functions [43].

其中是 GPS 和 IMU 时钟之间的时间偏移，边界位姿具有时间戳 ， , 是 矩阵的指数和对数函数[43]。

As evident from Eqs. (11)-(15), the GPS measurement model depends on both the IMU states and the GPS-IMU extrinsic and time offset, thus enabling online spatiotemporal GPS-IMU calibration. To update with this measurement in the MSCKF, we linearize it at the current estimate and have the following measurement Jacobians：

从方程式中可以看出。 (11)-(15)，GPS 测量模型依赖于 IMU 状态和 GPS-IMU 外在和时间偏移，从而实现在线时空 GPS-IMU 校准。为了在 MSCKF 中更新此测量值，我们在当前估计值处对其进行线性化，并具有以下测量雅可比行列式：





where is the skew symmetric matrix, and are left and right Jacobians of [43], and. With these, we are ready to perform EKF update. The details can be found in the companion technical report [44].

其中 是斜对称矩阵， 和 是 [43] 的左右雅可比矩阵，。有了这些，我们就可以进行 EKF 更新了。详细信息可以在配套的技术报告中找到[44]。

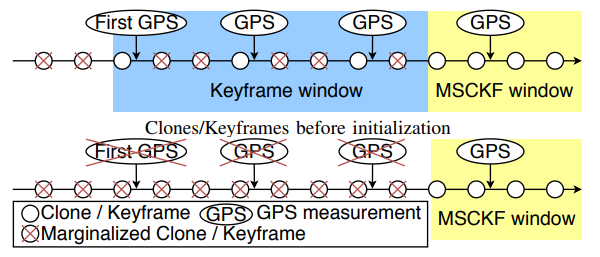


Fig. 3: The variation of windows during and after GPS-VIO initialization. GPS-VIO inserts keyframes after the first GPS measurement is received, and marginalize the keyframes after the initialization, leaving only the standard camera clones.

图 3：GPS-VIO 初始化期间和之后窗口的变化。 GPS-VIO 在接收到第一个 GPS 测量值后插入关键帧，并在初始化后边缘化关键帧，只留下标准相机克隆。

1. GPS-VIO INITIALIZATION

As mentioned earlier, when performing an EKF update with GPS measurements Eq. (11), the 4 d.o.f frame transformation must be known. To find this, we can collect two sets of position estimates of the GPS receiver in two different frames and formulate a non-linear optimization problem to align them. This process requires us to have estimates of the GPS receiver positions in both {E} and {V } frames. In the case of inaccurate GPS measurements, alignment using a short trajectory length may result in a poor transformation estimate due to the true trajectory being buried in the large measurement noise. As shown in simulations in Section VI, the smart use of longer trajectories allows for accurate alignment even with high noise

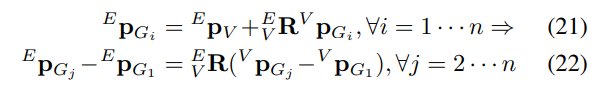
如前所述，当使用 GPS 测量执行 EKF 更新时，方程式。 (11)，必须知道 4 自由度框架变换 。为了找到这一点，我们可以在两个不同的坐标系中收集两组 GPS 接收器的位置估计，并制定一个非线性优化问题来对齐它们。这个过程要求我们在 {E} 和 {V} 系中估计 GPS 接收器的位置。在 GPS 测量不准确的情况下，使用短轨迹长度的对齐可能会导致较差的变换估计，因为真实轨迹被隐藏在大的测量噪声中。如第 VI 节中的模拟所示，巧妙地使用更长的轨迹，即使在高噪音的情况下也能实现精确对准

In the standard MSCKF-VIO, the current sliding window typically contains a very short and most recent portion of the trajectory, which does not support reliable GPS-VIO initialization. Therefore, we augment our state by selectively keeping the clone poses (i.e., keyframes) that bound GPS measurements at a fixed temporal frequency. As illustrated in Fig. 3, once we reach the desired trajectory length, we perform interpolation for all GPS measurement times that fall within the keyframe window to find the corresponding position estimates in the VIO frame.

在标准 MSCKF-VIO 中，当前滑动窗口通常包含轨迹的非常短且最近的部分，不支持可靠的 GPS-VIO 初始化。因此，我们通过选择性地保持以固定时间频率绑定 GPS 测量的克隆姿势（即关键帧）来增强我们的状态。如图 3 所示，一旦我们达到所需的轨迹长度，我们对落在关键帧窗口内的所有 GPS 测量时间执行插值，以在 VIO 坐标系中找到相应的位置估计。

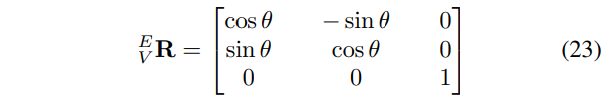
Given a set of GPS position measurements in the ENU frame within the keyframe window and the corresponding interpolated positions in the VIO frame, we use the following geometric constraints to derive the frame initialization:

给定关键帧窗口内 ENU 帧 中的一组 GPS 位置测量值和 VIO 帧 , 中的相应插值位置，我们使用以下几何结构导出帧初始化的应变：



As mentioned earlier, there is a 4 d.o.f (instead of 6 d.o.f) transformation including 3 d.o.f translation and 1 d.o.f for yaw between the ENU and VIO frames due to the fact that both frames are gravity aligned, which entails that we can simply use the rotation about the global z-axis with yaw angle θ:

如前所述，由于两个框架都是重力对齐的，因此在 ENU 和 VIO 框架之间存在 4 自由度（而不是 6 自由度）转换，包括 3 自由度平移和 1 自由度用于偏航，这意味着我们可以简单地使用以偏航角 θ 绕全局 z 轴旋转：



With (23) we can re-write (22) as the following linear constraint:

使用 (23) 我们可以将 (22) 重写为以下线性约束：



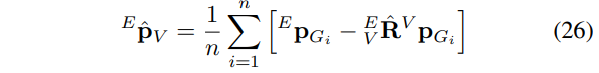
Stacking all these constraints yields the following linear least-squares with quadratic constraint, which can be solved for w, e.g., by Lagrangian multipliers [39]:

堆叠所有这些约束产生以下具有二次约束的线性最小二乘法，可以通过例如拉格朗日乘数 [39] 求解 w：



The solution of (25) immediately provides the sought rotation . We substitute it into (21) and solve for as:

(25) 的解立即提供了所寻求的旋转. 。我们将其代入 (21) 并将求解为：



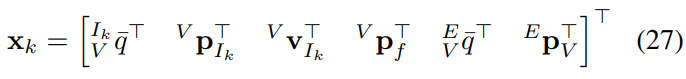
The resulting initial guess of the GPS-VIO frame transformation is further corrected using delayed initialization [42], [45], which appends the transform to the state in a probabilistic fashion. Specifically, by augmenting the state vector with the transformation along with an infinite covariance prior for these new variables, we perform the standard EKF update using all collected GPS measurements. After initialization, we marginalize all the keyframes to reduce the state to the original state size (see Fig. 3).

使用延迟初始化 [42]、[45] 进一步校正 GPS-VIO 帧变换的结果初始猜测，它以概率方式将变换附加到状态。具体来说，通过使用变换以及这些新变量的无限协方差先验来增加状态向量，我们使用所有收集的 GPS 测量值执行标准 EKF 更新。初始化后，我们边缘化所有关键帧以将状态减小到原始状态大小（见图 3）。

1. OBSERVABILITY ANALYSIS OF GPS-VIO

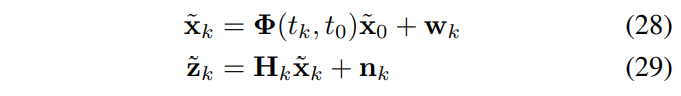
As system observability plays an important role state estimation [27], [46], in this section we perform an observability analysis for the proposed GPS-aided VIO system to gain insights about state/parameter identifiability. For concise presentation, we consider a simplified case where the state does not contain biases or stochastic clones and assumes a single feature with perfectly synchronized and calibrated sensors, while the results can be extended to general cases:

由于系统可观测性在状态估计 [27]、[46] 中起着重要作用，因此在本节中，我们对所提出的 GPS 辅助 VIO 系统进行观测能力分析，以深入了解状态/参数可识别性。为了简洁的介绍，我们考虑一个简化的情况，其中状态不包含偏差或随机克隆，并假设具有完美同步和校准传感器的单一特征，而结果可以扩展到一般情况：



The linearized error state evolution and residuals of both the GPS and visual measurement are generically given by (see (5), (7) and (11)):

GPS 和视觉测量的线性化误差状态演变和残差一般由（见（5）、（7）和（11））给出：



Given this linearized system, we have the following result:

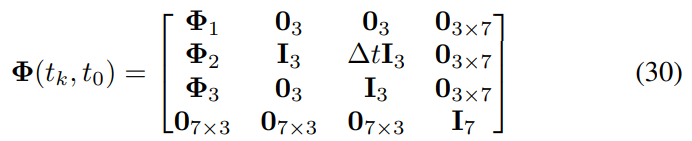
给定这个线性化系统，我们有以下结果：

Lemma 5.1: If estimating states in the VIO frame, even with global GPS measurements, the GPS-VIO system remains unobservable and has four unobservable directions.

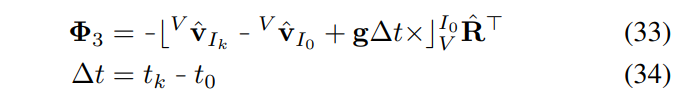
引理 5.1：如果在 VIO 框架中估计状态，即使使用全球 GPS 测量，GPS-VIO 系统仍然是不可观测的，并且有四个不可观测的方向。

Proof: We first compute the state transition matrix (6):

证明：我们首先计算状态转移矩阵（6）：

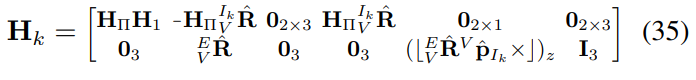


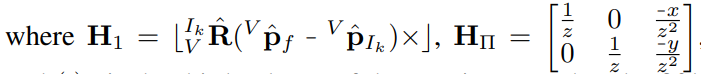
其中



Linearization of (7) and (11) yields the following measurement Jacobians:

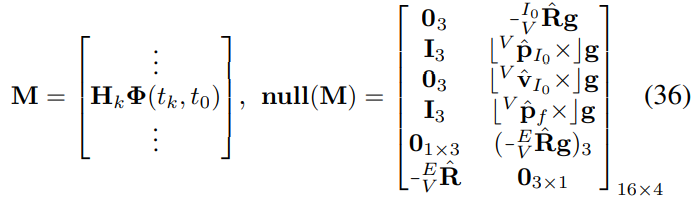
(7) 和 (11) 的线性化产生以下测量雅可比行列式





and is the third column of the matrix. Note that the fifth column is 5 by 1, because we have 1 d.o.f for the {E} to {V } rotation. Now we can construct the observability matrix M (see [47]) and compute its null space as:

(·)z 是矩阵的第三列。请注意，第五列是 5 x 1，因为 {E} 到 {V} 的旋转有 1 个自由度。现在我们可以构造可观察性矩阵 M（参见 [47]）并将其零空间计算为：



where is the third element of the vector. The span of the columns of this matrix encodes the unobservable subspace. By inspection, the first block column corresponds to the translation of {V } relative to {E} and the second block column to the rotation of {V } with respect to {E} along the axis of gravity. It thus becomes clear that the GPS-VIO system in the VIO frame has these four unobservable directions which are essentially inherited from the standard VIO [27], [28].

其中 (·)3 是向量的第三个元素。这个矩阵的列的跨度编码了不可观察的子空间。通过检查，第一个块列对应于 {V} 相对于 {E} 的平移，第二个块列对应于 {V} 相对于 {E} 沿重力轴的旋转。因此很明显，VIO 帧中的 GPS-VIO 系统具有这四个不可观测的方向，这些方向本质上是从标准 VIO [27]、[28] 继承的。

While the above results seem to be counter-intuitive given the availability of global GPS measurements, the root cause of this unobservability is the gauge freedom of the 4 d.o.f GPS-VIO frame transformation. Thus even though we utilize global measurements, the system maintains a non-trivial null space. Unobservable directions are known to cause inconsistency issues for linearized estimators as these null spaces falsely disappear due to numerical errors. Therefore the estimator gains information in spurious directions, hurting overall consistency and accuracy, unless special techniques are utilized [27], [48]. To address this issue, we perform state estimation directly in the ENU global frame of reference once initialized, which can be shown to make the system fully observable.

虽然鉴于全球 GPS 测量的可用性，上述结果似乎与直觉相反，但这种不可观测性的根本原因是 4 自由度 GPS-VIO 框架变换的规范自由度。因此，即使我们使用全局测量，系统也会保持一个非平凡的零空间。

众所周知，不可观察的方向会导致线性化估计器的不一致问题，因为这些零空间会由于数值错误而错误地消失。因此，估计器在虚假方向上获取信息，损害整体一致性和准确性，除非使用特殊技术 [27]、[48]。为了解决这个问题，

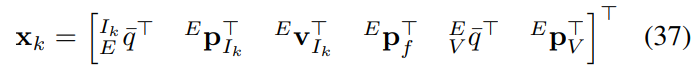
一旦初始化，我们就直接在 ENU 全局参考框架中执行状态估计，这可以证明系统完全可观察。

Lemma 5.2: If estimating states in the ENU frame, the GPS-VIO system is fully observable.

引理 5.2：如果估计 ENU 帧中的状态，则 GPS-VIO 系统是完全可观察的。

Proof: The simplified state in the ENU frame is:

证明：ENU帧中的简化状态为：

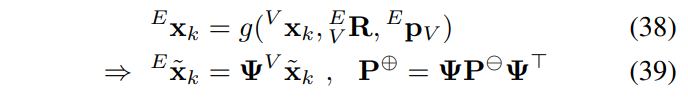


Then the state transition matrix of the new state is equivalent to Eq. (30) with all parameters that are in {V } are now in {E}. Also, the corresponding GPS measurement Jacobian is (see Eq. (11)). Clearly, the multiplication of with **null(M)** does not yield a zero matrix which means the four unobservable directions of Eq. (27) are now observable given GPS measurements. Since VIO is known to have four unobservable directions [27], [28], we can conclude that the state in the ENU, see Eq. (37), is fully observable

那么新状态 的状态转移矩阵等价于方程。 (30) {V} 中的所有参数现在都在 {E} 中。此外，对应的 GPS 测量值 Ja cobian 为 （参见方程（11））。显然， 具有 null(M) 的不会产生零矩阵，这意味着等式的四个不可观察的方向。 (27) 现在可以在 GPS 测量中观察到。由于已知 VIO 有四个不可观察的方向 [27]、[28]，我们可以得出结论，ENU 中的状态，参见方程式。 (37), 是完全可观察的

As a final remark about the proposed GPS-VIO estimator, based on the above lemma, after GPS-VIO initialization, we therefore transform the state from {V } to the {E} and propagate the error state and covariance based on the linearization of this transform function g(·) as follows:

作为对所提出的 GPS-VIO 估计器的最后评论，基于上述引理，在 GPS-VIO 初始化之后，我们因此将状态从 {V} 转换为 {E} 并基于线性化传播误差状态和协方差这个变换函数g(·)如下：



where **Ψ** is the Jacobian matrix [44]. We note that the {E} to {V } transformation inserted into the state during initialization, see Section IV, has been marginalized since all measurements can now be written directly in terms of the remaining state variables.

其中是雅可比矩阵[44]。我们注意到在初始化期间插入到状态中的 {E} 到 {V} 转换（参见第 IV 节）已被边缘化，因为现在可以直接根据剩余状态变量编写所有测量值。

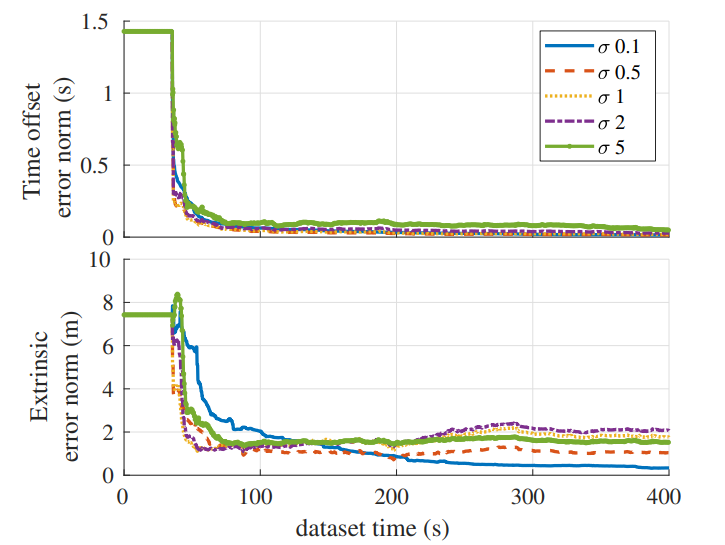
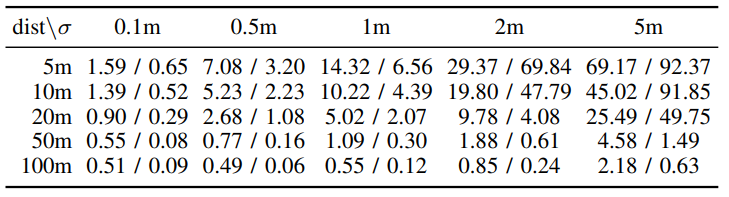


Fig. 4: The calibration errors respect to the size of GPS measurement noise.

图 4：相对于 GPS 测量噪声大小的校准误差。

TABLE I: Average position and orientation errors over ten runs for different initialization distances and GPS noise values in units of meters/degree.

表 I：十次运行中不同初始化距离和 GPS 噪声值的平均位置和方向误差，以米/度为单位。



1. SIMULATION RESULTS

The proposed GPS-VIO was implemented within OpenVINS [49] which provides both simulation and evaluation utilities. The key simulation parameters are: maximum of 15 clones, maximum of 100 actively tracked features with 1 pixel Gaussian noise, while the IMU was simulated using realistic noise from a real sensor. The camera was simulated at 5Hz, while the GPS sensor was simulated with a lower frequency of 2Hz with varying measurement noises ranging from 0.1m to 5m. As shown in Fig. 1, the trajectory of the dataset is 9.1km in length, following that of a planar vehicle motion with an average velocity of 9m/s. Except for the results in Section VI-C, Monte-Carlo simulation results are reported over 10 runs.

提议的 GPS-VIO 在 Open VINS [49] 中实现，它提供了模拟和评估实用程序。关键的模拟参数是：最多 15 个克隆，最多 100 个主动跟踪特征和 1 像素高斯噪声，而 IMU 使用来自真实传感器的真实噪声进行模拟。相机以 5Hz 模拟，而 GPS 传感器以 2Hz 的较低频率进行仿真，测量噪声范围从 0.1m 到 5m。如图 1 所示，数据集的轨迹长度为 9.1km，遵循平均速度为 9m/s 的平面车辆运动轨迹。除第 VI-C 节中的结果外，Monte-Carlo 模拟结果报告了 10 次运行。

1. Initialization with Different Hyper-parameters
2. 使用不同的超参数初始化

To gain insight into how the initialization procedure is affected by GPS measurement noise and trajectory length, we simulated 0.1, 0.5, 1, 2 and 5m GPS measurement noise and 5, 10, 20, 50 and 100m initialization distance thresholds. To prevent biasing these results to the initial section of this particular trajectory, it is split into non-overlapping segments for each distance threshold. The initialization procedure was independently performed on each segment and the resulting statistics on the accuracy of the initialized VIO to ENU transform are shown in Table. I.

为了深入了解初始化过程如何受到 GPS 测量噪声和轨迹长度的影响，我们模拟了 0.1、0.5、1、2 和 5m GPS 测量噪声以及 5、10、20、50 和 100m 初始化距离阈值。为了防止将这些结果偏向该特定轨迹的初始部分，将其分为非每个距离阈值的重叠段。初始化过程在每个段上独立执行，初始化 VIO 到 ENU 转换精度的统计结果如表所示, Table I。

In general, the initialization errors are smaller with a larger distance threshold and with smaller GPS noise. The results indicate that reasonable accuracy for this transformation can be achieved after 50m for most realistic levels of GPS measurement noise. In practice, these results can be used to determine the needed distance threshold for different sensor uncertainties

一般来说，距离阈值越大，GPS 噪声越小，初始化误差越小。结果表明，对于最现实的 GPS 测量噪声水平，这种变换的合理精度可以在 50m 后实现。在实践中，这些结果可用于确定不同传感器不确定性所需的距离阈值

1. Calibration with Different GPS Noise Levels

In order to study the calibration convergence of the proposed system, we performed extrinsic calibration and time offset between the GPS and IMU with poor initial guesses. The groundtruth and initial guess for the extrinsic were [2.00 3.00 1.00]> and [5.40 1.65 6.62]> meters, while for the time offset they were 0 and 91.3 seconds. The calibration results for the first 400 seconds are shown in Fig. 4, which clearly demonstrates that the time offset calibration converges to near zero. The final converged extrinsic calibration error follows that of the GPS measurement noise except in the 5m σ case. This shows that the convergence of the extrinsic is highly dependent on the measurement noise and whose final error is on the order of the GPS measurements.

为了研究所提出系统的校准收敛性，我们在初始猜测较差的 GPS 和 IMU 之间进行了外部校准和时间偏移。外在的真实值和初始猜测是 [2.00 3.00 1.00]> 和 [5.40 1.65 6.62]> 米，而时间偏移量是 0 和 91.3 秒。前 400 秒的校准结果如图 4 所示，清楚地表明时间偏移校准收敛到接近零。除了 5m σ 的情况外，最终收敛的外部校准误差遵循 GPS 测量噪声的误差。这表明外在的收敛高度依赖于测量噪声，其最终误差在 GPS 测量的数量级上。

A representative run is shown in Fig. 5, all calibration was able to converge within the first 100 seconds of the dataset while remaining consistent. The static lines at the beginning of the each are from before initialization in the ENU and thus no GPS measurements that are required to update these parameters have been used. As expected, the y-error, which is mostly aligned with gravity in this scenario, shows little decrease in state uncertainty due to this axis corresponding to the normal of the plane of motion [21].

代表性运行如图 5 所示，所有校准都能够在数据集的前 100 秒内收敛，同时保持一致。每个开头的静态线来自 ENU 中的初始化之前，因此没有使用更新这些参数所需的 GPS 测量。正如预期的那样，在这种情况下，y 误差主要与重力对齐，由于该轴对应于运动平面的法线 [21]，因此状态不确定性几乎没有减少。

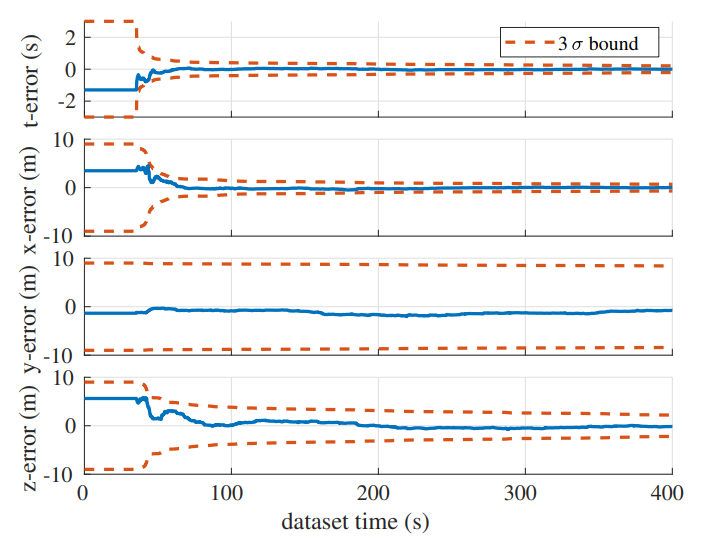


Fig. 5: Calibration errors of the proposed GPS-VIO with 5m GPS sensor noise and the corresponding consistency bounds. The blue lines are the errors and the red dotted lines are the 3 standard deviation bounds of each error

图 5：建议的 GPS-VIO 与 5m GPS 传感器噪声的校准误差和相应的一致性界限。蓝线是错误，红色虚线是每个错误的 3 个标准偏差范围

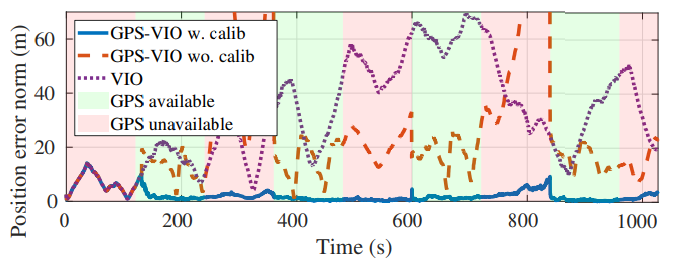


Fig. 6: The position error of each method in the case of intermittent GPS.

图 6：间歇 GPS 情况下各方法的位置误差。

1. Robustness to Intermittent GPS Signals

C. 对间歇性 GPS 信号的鲁棒性

To validate the robustness to intermittent GPS measurements, we simulated a series of GPS dropouts during which the GPS-VIO purely relied on visual and inertial information. The GPS dropouts lasted 120 seconds in length, the ENU to VIO transformation was set to identity to allow for comparison to pure VIO, and the calibration was perturbed as in the last simulation.

为了验证间歇性 GPS 测量的稳健性，我们模拟了一系列 GPS 丢失，在此期间 GPS-VIO 完全依赖视觉和惯性信息。 GPS 丢失持续了 120 秒，ENU 到 VIO 的转换被设置为身份，以便与纯 VIO 进行比较，并且校准像上次模拟一样受到干扰。

Fig. 6 shows the errors of GPS-VIO with online calibration, GPS-VIO without online calibration, and pure VIO. Before initialization in ENU all systems are purely VIO. After the initialization in the ENU at around 150 seconds, the proposed method begins fusing GPS data and thus quickly bounds its errors. It is clear that poor calibration can hurt the system’s estimation even in the presence of global measurements. Finally, the relatively good accuracy of the VIO is able to “bridge the gap” between the GPS-available regions, providing high-quality navigation estimates over the entire trajectory. Note that reducing error of the pure VIO at time around 800 seconds is because the trajectory has loops, which brings the estimation and groundtruth close by chance.

图 6 显示了带在线校准的 GPS-VIO、不带在线校准的 GPS-VIO 和纯 VIO 的误差。在 ENU 中初始化之前，所有系统都是纯 VIO。在 ENU 中大约 150 秒初始化后，所提出的方法开始融合 GPS 数据，从而快速限制其错误。

很明显，即使在存在全局测量的情况下，糟糕的校准也会损害系统的估计。最后，VIO 相对较好的精度能够“弥合”GPS 可用区域之间的差距，在整个轨迹上提供高质量的导航估计。

请注意，减少纯 VIO 在 800 秒左右的误差是因为轨迹有循环，这使得估计和 groundtruth 偶然接近。

VII. EXPERIMENTAL RESULTS

We further evaluate the proposed GPS-VIO in a realworld scenario. The trajectory begins in an indoor parking garage during which the system does not have access to GPS measurements until a minute in when it exits the structure. During the outdoor segment the vehicle travels several kilometers before returning to the same GPS-denied structure. The total length of the data is about 4.9km, and we used a monocular camera-imu pair, alongside two GPS receivers all of which were mounted rearward on the trunk of the collection vehicle. One low-cost GPS sensor was used for GPS measurements while the second provided RTK data for groundtruth. The covariance of each GPS measurement was computed by RTKLIB library [50]. We compared our system against the open sourced VINS-Fusion [12] system. We used 30 clones and a max of 150 features for the proposed GPSVIO while for VINS-Fusion a max of 200 features and a max solver time 0.04 were used. We note that VINS-Fusion does not take into account the GPS-IMU calibration.

我们进一步评估了在现实世界场景中提出的 GPS-VIO。轨迹从室内停车场开始，在此期间系统直到离开建筑物的一分钟后才能访问 GPS 测量值。在室外部分，车辆行驶了几公里，然后返回到相同的 GPS 结构。数据的总长度约为 4.9 公里，我们使用了单目相机-imu 对，以及两个 GPS 接收器，所有这些接收器都安装在收集车的后备箱上。一个低成本 GPS 传感器用于 GPS 测量，而第二个提供 RTK 数据用于地面实况。每个 GPS 测量的协方差由 RTKLIB 库 [50] 计算。我们将我们的系统与开源的 VINS-Fusion [12] 系统进行了比较。我们为建议的 GPS VIO 使用了 30 个克隆和最多 150 个特征，而对于 VINS-Fusion，使用了最多 200 个特征和最大求解器时间 0.04。我们注意到VINS-Fusion 不考虑 GPS-IMU 校准

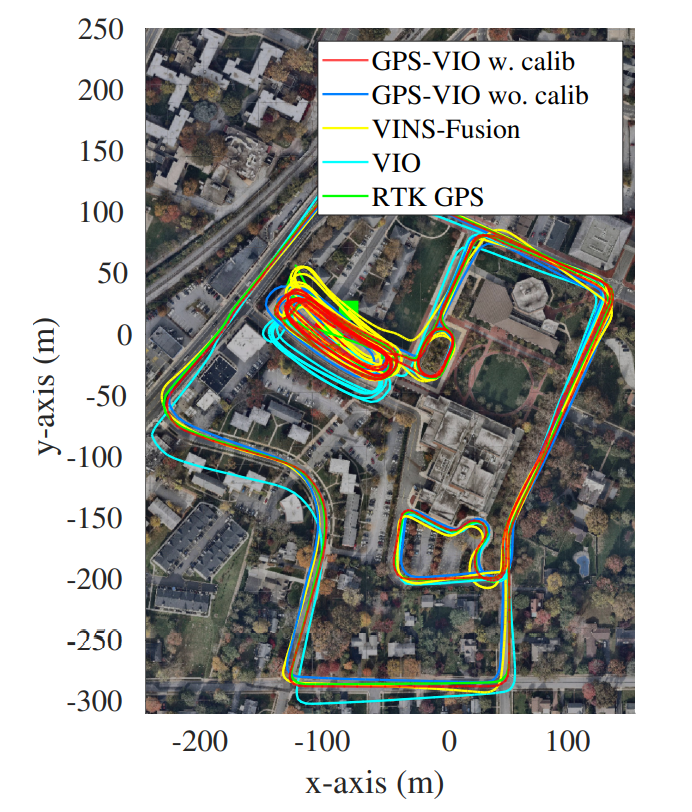


Fig. 7: The red line is GPS-VIO with online calibration, blue is GPS-VIO without online calibration, yellow is VINS-Fusion, light blue is VIO only, and green is RTK GPS. The green and red boxes denote the start end points of the trajectory.

图7：红线为在线校准的GPS-VIO，蓝色为未在线校准的GPS-VIO，黄色为VINS-Fusion，浅蓝色为仅VIO，绿色为RTK GPS。绿色和红色框表示轨迹的起点。

Fig. 7 shows the result of the experiment. The RMSE of each trajectory compared with RTK groundtruth whenever it is available are: 3.57m (GPS-VIO w. calib), 7.03m (GPS-VIO wo. calib), 6.66m (VINS-Fusion) and 15.95m (VIO). We gave the initial hand measured extrinsic value meters and time offset of zero for both GPS-VIO w. calib and GPS-VIO wo. calib. The extrinsic calibration converged to a value of at the end of the dataset which is within the expected convergence bounds associated with the 1-10 meters observed GPS noise. We found time offset calibration quickly converged to a nontrivial value of 90.85 seconds which given the 7.4m/s average vehicle velocity equates to 6.3m position error if not properly calibrated, and thus validates the need for online estimation of this parameter. As compared to VINS-Fusion our method can achieve higher accuracy while also being a light-weight single threaded estimator which runs in real time. We also note that since VINS-Fusion does not estimate VIO to ENU transform explicitly, its pose output may not suitable for real time applications, and Fig. 7 shows the final optimized trajectory after completion of the dataset.

图 7 显示了实验结果。与 RTK groundtruth 相比，每个轨迹的 RMSE 在可用时为：3.57m（GPS-VIO w. calib）、7.03m（GPS-VIO wo.calib）、6.66m（VINS-Fusion）和 15.95m（VIO ）。对于 GPS-VIO w，我们给出了初始手测外在值 [0.06 0.11 -0.03]^T 米和零时间偏移。calib 和 GPS-VIO wo。口径。外部校准在数据集末尾收敛到 [-0.49 0.70 -0.07]^T 的值，该值在与观测到的 1-10 米 GPS 噪声相关的预期收敛范围内。我们发现时间偏移校准很快收敛到 90.85 秒的非平凡值，给定 7。

如果没有正确校准，4m/s 的平均车辆速度相当于 6.3m 的位置误差，因此验证了该参数在线估计的必要性。与 VINS-Fusion 相比，我们的方法可以实现更高的精度，同时也是一种实时运行的轻量级单线程估计器。我们还注意到，由于 VINS-

Fusion 没有明确估计 VIO 到 ENU 的变换，它的位姿输出可能不适合实时应用，图 7 显示了数据集完成后的最终优化轨迹。

1. CONCLUSIONS AND FUTURE WORK
2. 结论和未来工作

In this paper, we have developed an efficient and robust GPS-VIO system that fuses GPS, IMU, and camera measurements in a tightly-coupled estimator. In particular, to robustify the system, we have focused on the online GPS-VIO spatiotemporal sensor calibration and frame initialization. The observability analysis shows that if estimating states naively in the VIO frame, the system remains unobservable as the standard VIO; however, this can be mitigated by transforming the system to the global ENU frame after GPSVIO frame initialization, which is exploited in the proposed GPS-VIO estimator. This system has been validated in both Monte-Carlo simulations and real-world experiments. In the future, we will integrate mapping capability into this GPSVIO system

在本文中，我们开发了一种高效且强大的 GPS-VIO 系统，该系统将 GPS、IMU 和相机测量值融合在一个紧密耦合的估计器中。特别是，为了增强系统的鲁棒性，我们专注于在线 GPS-VIO 时空传感器校准和帧初始化。可观察性分析表明，如果在 VIO 框架中天真地估计状态，则系统仍然像标准 VIO 一样不可观察；然而，这可以通过在 GPS VIO 帧初始化后将系统转换为全局 ENU 帧来缓解，这在所提出的 GPS-VIO 估计器中得到了利用。该系统已在蒙特卡罗模拟和实际实验中得到验证。未来，我们会将地图功能集成到这个 GPS VIO 系统中

REFERENCES

参考 [Cānkǎo]

[1] G. Huang, “Visual-inertial navigation: A concise review,” in Proc. International Conference on Robotics and Automation, Montreal, Canada, May 2019, pp. 9572–9582.

[2] T. Qin, P. Li, and S. Shen, “Vins-mono: A robust and versatile monocular visual-inertial state estimator,” IEEE Transactions on Robotics, vol. 34, no. 4, pp. 1004–1020, 2018.

[3] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, “On-manifold preintegration for real-time visual–inertial odometry,” IEEE Transactions on Robotics, vol. 33, no. 1, pp. 1–21, 2016. [4] G. Loianno, C. Brunner, G. McGrath, and V. Kumar, “Estimation, control, and planning for aggressive flight with a small quadrotor with a single camera and imu,” IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 404–411, 2016.

[5] G. Huang, M. Kaess, and J. Leonard, “Towards consistent visualinertial navigation,” in Proc. of the IEEE International Conference on Robotics and Automation, Hong Kong, China, May 31-June 7 2014, pp. 4926–4933.

[6] R. Mur-Artal and J. D. Tardos, “Orb-slam2: An open-source slam ´ system for monocular, stereo, and rgb-d cameras,” IEEE Transactions on Robotics, vol. 33, no. 5, pp. 1255–1262, 2017.

[7] P. Geneva, K. Eckenhoff, and G. Huang, “A linear-complexity EKF for visual-inertial navigation with loop closures,” in Proc. International Conference on Robotics and Automation, Montreal, Canada, May 2019, pp. 3535–3541.

[8] A. Almagbile, J. Wang, and W. Ding, “Evaluating the performances of adaptive kalman filter methods in gps/ins integration,” Journal of Global Positioning Systems, vol. 9, no. 1, pp. 33–40, 2010.

[9] C. Eling, L. Klingbeil, and H. Kuhlmann, “Real-time single-frequency gps/mems-imu attitude determination of lightweight uavs,” Sensors, vol. 15, no. 10, pp. 26 212–26 235, 2015. [10] R. Sun, K. Han, J. Hu, Y. Wang, M. Hu, and W. Y. Ochieng, “Integrated solution for anomalous driving detection based on beidou/gps/imu measurements,” Transportation research part C: emerging technologies, vol. 69, pp. 193–207, 2016.

[11] T. Zhang and X. Xu, “A new method of seamless land navigation for gps/ins integrated system,” Measurement, vol. 45, no. 4, pp. 691–701, 2012.

[12] T. Qin, S. Cao, J. Pan, and S. Shen, “A general optimization-based framework for global pose estimation with multiple sensors,” arXiv preprint arXiv:1901.03642, 2019.

[13] Y. Lee, J. Yoon, H. Yang, C. Kim, and D. Lee, “Camera-gpsimu sensor fusion for autonomous flying,” in Eighth International Conference on Ubiquitous and Future Networks (ICUFN), 2016, pp. 85–88.

[14] C. V. Angelino, V. R. Baraniello, and L. Cicala, “Uav position and attitude estimation using imu, gnss and camera,” in 15th International Conference on Information Fusion, 2012, pp. 735–742.

[15] T. Zsedrovits, P. Bauer, A. Zarandy, B. Vanek, J. Bokor, and T. Roska, “Error analysis of algorithms for camera rotation calculation in gps/imu/camera fusion for uav sense and avoid systems,” in International Conference on Unmanned Aircraft Systems (ICUAS), 2014, pp. 864–875.

[16] T. Oskiper, S. Samarasekera, and R. Kumar, “Multi-sensor navigation algorithm using monocular camera, imu and gps for large scale augmented reality,” in IEEE international symposium on mixed and augmented reality (ISMAR), 2012, pp. 71–80.

[17] S. Shen, Y. Mulgaonkar, N. Michael, and V. Kumar, “Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft mav,” in IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 4974–4981.

[18] F. M. Mirzaei and S. I. Roumeliotis, “A kalman filter-based algorithm for imu-camera calibration: Observability analysis and performance evaluation,” IEEE transactions on robotics, vol. 24, no. 5, pp. 1143– 1156, 2008.

[19] P. Furgale, J. Rehder, and R. Siegwart, “Unified temporal and spatial calibration for multi-sensor systems,” in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013, pp. 1280–1286.

[20] M. Li and A. I. Mourikis, “Online temporal calibration for camera– imu systems: Theory and algorithms,” The International Journal of Robotics Research, vol. 33, no. 7, pp. 947–964, 2014.

[21] Y. Yang, P. Geneva, K. Eckenhoff, and G. Huang, “Degenerate motion analysis for aided INS with online spatial and temporal calibration,” IEEE Robotics and Automation Letters (RA-L), vol. 4, no. 2, pp. 2070– 2077, 2019.

[22] H. Wegmann, “Image orientation by combined (a) at with gps and imu,” International Archives Of Photogrammetry Remote Sensing And Spatial Information Sciences, vol. 34, no. 1, pp. 278–283, 2002.

[23] M. Bryson, M. Johnson-Roberson, and S. Sukkarieh, “Airborne smoothing and mapping using vision and inertial sensors,” in IEEE International Conference on Robotics and Automation, 2009, pp. 2037–2042.

[24] X. Kong, E. M. Nebot, and H. Durrant-Whyte, “Development of a nonlinear psi-angle model for large misalignment errors and its application in ins alignment and calibration,” in Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C), vol. 2, 1999, pp. 1430–1435.

[25] K. Hausman, J. Preiss, G. S. Sukhatme, and S. Weiss, “Observabilityaware trajectory optimization for self-calibration with application to uavs,” IEEE Robotics and Automation Letters, vol. 2, no. 3, pp. 1770– 1777, 2017.

[26] A. Ramanandan, M. Chari, and A. Joshi, “Systems and methods for using a global positioning system velocity in visual-inertial odometry,” Aug. 6 2019, uS Patent 10,371,530. [27] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, “Observability-constrained vision-aided inertial navigation,” University of Minnesota, Dept. of Comp. Sci. & Eng., MARS Lab, Tech. Rep, vol. 1, p. 6, 2012.

[28] J. Kelly and G. S. Sukhatme, “Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration,” The International Journal of Robotics Research, vol. 30, no. 1, pp. 56–79, 2011.

[29] B. K. Horn, “Closed-form solution of absolute orientation using unit quaternions,” Josa a, vol. 4, no. 4, pp. 629–642, 1987.

[30] D. P. Shepard and T. E. Humphreys, “High-precision globallyreferenced position and attitude via a fusion of visual slam, carrierphase-based gps, and inertial measurements,” in IEEE/ION Position, Location and Navigation Symposium-PLANS, 2014, pp. 1309–1328.

[31] S. Umeyama, “Least-squares estimation of transformation parameters between two point patterns,” IEEE Transactions on Pattern Analysis & Machine Intelligence, no. 4, pp. 376–380, 1991.

[32] I. Zendjebil, F. Ababsa, J.-Y. Didier, and M. Mallem, “A gps-imucamera modelization and calibration for 3d localization dedicated to outdoor mobile applications,” in ICCAS 2010, 2010, pp. 1580–1585.

[33] Z. Zhang and D. Scaramuzza, “A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry,” in 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018, pp. 7244–7251.

[34] A. Waegli and J. Skaloud, “Optimization of two gps/mems-imu integration strategies with application to sports,” GPS solutions, vol. 13, no. 4, pp. 315–326, 2009.

[35] B. Khaleghi, A. El-Ghazal, A. R. Hilal, J. Toonstra, W. B. Miners, and O. A. Basir, “Opportunistic calibration of smartphone orientation in a vehicle,” in IEEE 16th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2015, pp. 1– 6.

[36] S. A. Shaukat, K. Munawar, M. Arif, A. I. Bhatti, U. I. Bhatti, and U. M. Al-Saggaf, “Robust vehicle localization with gps dropouts,” in 6th International Conference on Intelligent and Advanced Systems (ICIAS), 2016, pp. 1–6.

[37] J. Almazan, L. M. Bergasa, J. J. Yebes, R. Barea, and R. Arroyo, ´ “Full auto-calibration of a smartphone on board a vehicle using imu and gps embedded sensors,” in IEEE Intelligent Vehicles Symposium (IV), 2013, pp. 1374–1380.

[38] A. I. Mourikis and S. I. Roumeliotis, “A multi-state constraint kalman filter for vision-aided inertial navigation,” in Proceedings 2007 IEEE International Conference on Robotics and Automation, 2007, pp. 3565–3572.

[39] C. X. Guo, K. Sartipi, R. C. DuToit, G. A. Georgiou, R. Li, J. OLeary, E. D. Nerurkar, J. A. Hesch, and S. I. Roumeliotis, “Resource-aware large-scale cooperative three-dimensional mapping using multiple mobile devices,” IEEE Transactions on Robotics, vol. 34, no. 5, pp. 1349– 1369, 2018.

[40] N. Trawny and S. I. Roumeliotis, “Indirect Kalman filter for 3D attitude estimation,” University of Minnesota, Dept. of Comp. Sci. & Eng., Tech. Rep., Mar. 2005.

[41] C. Hertzberg, R. Wagner, U. Frese, and L. Schroder, “Integrating ¨ generic sensor fusion algorithms with sound state representations through encapsulation of manifolds,” Information Fusion, vol. 14, no. 1, pp. 57–77, 2013.

[42] M. Li, “Visual-inertial odometry on resource-constrained systems,” Ph.D. dissertation, UC Riverside, 2014

[43] G. Chirikjian, Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications. Springer Science & Business Media, 2011, vol. 2.

[44] W. Lee, K. Eckenhoff, P. Geneva, and G. Huang, “Gps-aided visualinertial navigation in large-scale environments,” Robot Perception and Navigation Group (RPNG), University of Delaware, Tech. Rep., 2019. [Online]. Available: http://udel.edu/∼ghuang/papers/tr gps-vio.pdf

[45] R. C. DuToit, J. A. Hesch, E. D. Nerurkar, and S. I. Roumeliotis, “Consistent map-based 3d localization on mobile devices,” in 2017 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017, pp. 6253–6260.

[46] G. Huang, “Improving the consistency of nonlinear estimators: Analysis, algorithms, and applications,” Ph.D. dissertation, Department of Computer Science and Engineering, University of Minnesota, 2012. [Online]. Available: https://conservancy.umn.edu/handle/11299/ 146717

[47] Z. Chen, K. Jiang, and J. Hung, “Local observability matrix and its application to observability analyses,” in Proc. of the 16th Annual Conference of IEEE, Pacific Grove, CA, Nov. 27–30, 1990, pp. 100– 103.

[48] G. Huang, A. I. Mourikis, and S. I. Roumeliotis, “A first-estimates Jacobian EKF for improving SLAM consistency,” in Proc. of the 11th International Symposium on Experimental Robotics, Athens, Greece, July 14–17, 2008.

[49] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, “Openvins: A research platform for visual-inertial estimation,” in Proc. of the IEEE International Conference on Robotics and Automation, Paris, France, 2020. [Online]. Available: https://github.com/rpng/open vins

[50] T.Takasu, “Rtklib: An open source program package for gnss positioning.” [Online]. Available: https://http://www.rtklib.com